

## Description

# *STRESSED SEMICONDUCTOR DEVICE STRUCTURES HAVING GRANULAR SEMICONDUCTOR MATERIAL*

### BACKGROUND OF INVENTION

- [0001] The present invention relates to semiconductor device structures, such as a CMOS device structure including both nFET and pFET devices
- [0002] Mobility enhancements are important to future semiconductor, e.g. CMOS device technologies. Performance improvements from conventional process technologies are becoming extremely difficult to achieve. Methods to stress Si channels include: using SiGe which imparts stress from the bottom of the channel; different shallow trench isolation (STI) material choices which impart stresses from various sides, and SiN etch stop layers which also impart longitudinal stress from the sides. Drawbacks from the SiGe buffer layer or implanted-anneal-buffer approach with a strained Si cap layer are well known. Drawbacks include

dislocations that impact yield severely, along with significant difficulty controlling As diffusion enhancements. Further, the process is quite complicated and costly. The STI approach is less costly but is not self-aligned to the gate and has RX size sensitivity. The less costly approach of using nitride etch stop layers to create stress does produce some benefit, but the benefit is believed to be relatively marginal.

#### **SUMMARY OF INVENTION**

[0003] The present invention improves device performance using channel mobility enhancement. The present invention improves mobility from the top of the channel by using the stress properties of properly modulated polysilicon gate stacks. Prior to the present invention these stress properties were very difficult to control. However, the present invention includes a method and structure which use small grain polysilicon to control stress properties. The present invention provides a method and a structure to impart compressive stress to the pFET channel and a tensile stress to the nFET channel. Other embodiments include imparting compressive stress to the pFET channel while preventing compressive stress from being imparted to the nFET channel. Another embodiment of this invention in-

cludes imparting tensile stress to the nFET channel while preventing the tensile stress from being imparted to the pFET channel.

[0004] The method of the present invention disposes (e.g., deposits) small grain poly Si films in gate electrode stacks, which films become highly tensile as a result of annealing. The tensile film stress can be suppressed by annealing with a low temperature deposited SiN film disposed on top of the poly Si film.

[0005] It is a principal object of the present invention to enhance channel mobility in semiconductor devices.

[0006] According to the present invention, a method of fabricating a semiconductor device structure includes providing a substrate; providing an electrode on the substrate; forming a recess in the electrode, the recess having an opening; disposing a small grain semiconductor material within the recess; covering the opening to contain the small grain semiconductor material within the recess, and then annealing the resultant structure. The present invention also includes a semiconductor device structure.

[0007] Further and still other objects of the present invention will become more readily apparent in light of the following detailed description when taken in conjunction with the

accompanying drawing figures.

## **BRIEF DESCRIPTION OF DRAWINGS**

[0008] Figures 1–9 are side schematic views of process steps according to a preferred embodiment of the present invention.

[0009] Figure 10 is a diagram showing simulated stress contours in the gate and channel of the preferred inventive structure.

## **DETAILED DESCRIPTION**

[0010] Turning now to the figures, and Fig. 1 in particular, a semiconductor substrate 1 is provided. The semiconductor substrate is a bulk Si substrate, an SOI substrate, or a stressed (strained) Si substrate. Alternatively, the substrate is a hybrid substrate which includes more than one surface orientation. The substrate alternatively includes a semiconductor material other than Si, such as Ge or any combination of Group III–V elements or Group II–V elements.

[0011] After an initial substrate cleaning procedure (conventional), an isolation scheme is carried out. As is well known in semiconductor manufacturing, the isolation scheme is used to separate selected devices electrically

from each other. The isolation scheme may be a standard or a modified shallow trench isolation (STI) scheme. The STI 2 is shown in Fig. 1. Alternatively, the isolation is accomplished using a LOCOS process or mesa isolation scheme, as is well known in the art of fabricating semiconductor devices. For various known or conventional processes for fabricating semiconductor devices, see VLSI Technology, 2<sup>nd</sup> Edition, by S.M. Sze, (McGraw Hill Publishing Co., 1988).

[0012] After isolation 2 is formed, a conventional gate oxide pre-cleaning process is performed. As is the case in high performance logic fabrication processes, various conventional gate oxide processes may be used to fabricate devices having different gate oxide thicknesses. The gate oxide 3 is formed, for example, using a conventional thermal oxidation process. The oxide 3 is formed using  $N_2O$ ,  $NO$ ,  $O_2$  or any combination of them. The oxide may be nitridized using a conventional plasma process. Alternatively, the gate oxide may be formed using a base oxide followed by deposition of a high k gate dielectric such as aluminum oxide, or hafnium oxide, or another high k gate dielectric. The gate dielectric material 3 has an approximately (+ 10%) uniform thickness in the range of about (+ 10%)

0.6nm to about 7nm.

[0013] Next, a film 4 is deposited over the entire wafer structure 1, 2, 3. The film 4 is used as a disposable (removable) or semi-disposable gate electrode material. In a preferred embodiment, the film 4 includes a polysilicon (poly Si) material having an approximately uniform thickness or height (T) in the range of about 80nm to about 150nm. A deposition technique such as low pressure chemical vapor deposition (LPCVD) or rapid thermal chemical vapor deposition (RPCVD) is used to deposit the removable gate electrode material 4. The resulting structure is shown in Fig. 1. The poly Si layer 4 preferably has a standard grain size in a range of about one nm to about 40nm.

[0014] Next, a conventional lithography process is used to pattern photoresist images on the top of the removable gate electrode material 4. The photoresist images, not shown in the figures, are used to transfer desired features into the removable gate electrode material 4 by using a conventional dry etching process. The dry etching process includes or several chemistries capable of etching the removable gate electrode material 4 selectively with respect to the gate oxide material 3. The structure shown in Fig. 2 shows the fully patterned removable gate electrode 5 for a

nFET gate stack 3,5 and a removable gate electrode 6 for a pFET gate stack 3,6.

[0015] A conventional gate reoxidation process, not shown in the figures, is then used, as is commonly done in high performance logic manufacturing processes. The reox is formed by using a thermal oxidation process to achieve an approximately uniform thickness from about one nm to about 7nm. Following the reox process, a block mask is patterned over the pFET regions, using a conventional photolithography process. The block (e.g., resist) mask(not shown in the figures) is used to block or prevent the pFET regions from being implanted, while the appropriate nFET regions are being implanted. The nFET extensions 7 and halos (not shown) are implanted using a low energy As and B implant, respectively. The resist mask is then removed using a dry or wet process. Another block mask (not shown) is patterned over the nFET regions. The pFET extensions 8 and halos (not shown) are implanted using a low energy  $\text{BF}_2$  or B implant and As implant, respectively. The extension implant profiles 7,8 for the nFET and the pFET are shown in Fig. 2.

[0016] After the extension and halo implantation, a dielectric liner layer 9 is formed over the entire wafer structure (Fig.

3). The dielectric film to be used as the liner layer 9 is preferably SiN which is deposited by CVD or RTCVD or any other suitable deposition technique. A purpose of the liner layer 9 is to provide a CMP stop layer for the next process. Additionally, the SiN liner 9 will be etched, at a later point in the process flow, to form a source-drain spacer. The resulting structure is shown in Fig. 3.

[0017] The next step in the process flow is to deposit an oxide film 10. The oxide film 10 is deposited and planarized using chemical mechanical polishing (CMP). The film 10 is deposited using, for example a high density plasma (HDP) process. The top of the liner 9 over the removable gate electrode 5 is removed by using a dry etching process that is capable of etching silicon nitride but does not etch appreciable amounts of oxide or poly Si. The planarized oxide film 10 and SiN liner 9 structure is shown in Fig. 4, after the top portion of the liner layer 9 has been removed.

[0018] An important aspect of this invention is that, at this point in the inventive process flow, a gate recess process is used to remove the poly Si, either completely or partially, from the gate electrodes 5, 6. A preferred embodiment in which the poly Si is partially removed is shown in Fig. 5. The poly Si is recessed using any suitable dry or wet etch



process. A portion 12 and a portion 13 of the original poly Si is left remaining, and has an approximately uniform thickness in a range of about one nm to about 20nm. A recessed portion 12 of the nFET gate electrode and a recessed portion 13 of the pFET gate electrode are shown in Fig. 5. In another embodiment (not shown), the poly Si is completely removed. If this embodiment is used, then a conventional gate oxide pre-clean process followed by a conventional gate oxidation process is next performed.

[0019] Another important aspect of this invention is that, following the gate recess process, a small grain poly Si is deposited over the entire wafer. Small grain polysilicon is known from Shimizu, S. et al. Proceedings of the 1997 Symposium on VLSI Technology, Kyoto, Japan 10–12 June 1997, and also from Silicon Processing for The VLSI Era, Vol 1– Process Technology, by S. Wolf, 1999. The grain structure preferably is in a range of about one nm to about 50nm. A more preferable grain size is a substantially uniform size in a range of about five nm to 30nm. The polySi is deposited by RTCVD or LPCVD. Next, the polySi is planarized and recessed from the top of the oxide layer 10 using, for example, CMP and a dry etch. Both CMP and dry etch processes are capable of removing the poly Si selec-

tively relative to the  $\text{SiO}_2$  layer 10. After CMP and dry etch, the inventive gate electrode structure 14, 12 for the nFET and the inventive gate electrode structure 13, 18 for the pFET are shown in Fig. 6.

[0020] At this point in the inventive method, a hard mask 15 shown in Fig. 7 is patterned over the nFET region. The mask 15 is patterned using a dielectric film including SiN and a conventional photolithography procedure. The hard mask 15 is deposited using a low temperature deposition process such as plasma enhanced chemical vapor deposition with a temperature in the range of about  $350^\circ\text{C}$  to about  $700^\circ\text{C}$ . Preferably, the material which forms the hard mark 15 is deposited at a temperature below  $550^\circ\text{C}$  or lower, to prevent re-growth of the poly Si grains. The whole structure is next subjected to an anneal using a temperature from about  $500^\circ\text{C}$  to about  $600^\circ\text{C}$  for about one hour. Alternatively, a rapid thermal anneal may be carried out at about  $700^\circ\text{C}$  to about  $1000^\circ\text{C}$  for about one second to about seven seconds. In another embodiment, the structure 12, 14 is annealed at a future point in the process. The small grain poly Si remains small, in the range preferably of about five nm to about 30nm for the nFET gate electrode portion 17 because the SiN cap layer

15 is present. However, the grains in the pFET gate electrode portion 18 grow significantly to grain sizes of greater than about 30nm.

[0021] The present inventors believe that the inventive process (e.g., with respect to Fig. 7) represents a notable departure from conventional process technology. If the small grain poly Si is subjected to the standard thermal budget, then the poly Si grain growth causes a severe increase in tensile stress. The tensile stress creates a compressive stress in the channel region which degrades electron mobility and limits the performance for the nFET. See the simulations of Fig.10. The inventors believe that the grain growth and resulting severe increase in tensile stress is almost completely eliminated by annealing with the SiN hard mask 15 disposed over the nFET region. The grain structure can be optimized for each device independently by depositing the small grain poly Si into the recess and annealing the nFET with the SiN hard mask. This step appears to result in a notable improvement in nFET device performance.

[0022] The next step in the inventive process is to remove the SiN hard mask 15 from the entire horizontal part of the oxide film 10 or from the entire horizontal part (except over the portion 17, as shown in Fig.8). Because the poly Si

was recessed as described previously, a portion 19 of the hard mask 15 is left remaining (disposed) in the recessed hole as shown in Fig. 8. The purpose of the structure 19 is to prevent the grains from growing during subsequent conventional thermal cycles common in standard state-of-the art semiconductor manufacturing process technology. After the SiN etch, the oxide film 10 is removed using a suitable dry or wet etch process capable of removing the  $\text{SiO}_2$  film selectively relative to the SiN and the poly Si materials. A dry directional etch process is next performed on the liner layer 9 to form the spacers 20, 25 as shown in Fig.8. Although not essential for the present invention, the spacers 20 of the pFET and the spacers 25 of the nFET may have differing heights as shown. A similar block mask and implant process as was used to form the nFET and pFET extension regions 7, 8 is used to form the nFET source-drain regions 21 and the pFET source-drain region 22 shown in Fig. 8. A rapid thermal anneal is next performed to activate the junctions. Because the nFET continues to have the SiN layer 19 present, grain growth in the nFET gate is suppressed, thereby minimizing the tension in the gate electrode stack 3, 12, 17 and subsequent compression in the channel region below the stack.

[0023] The remaining portion 19 over the gate electrode stack is next removed using a wet or dry etch process. Next, a silicide pre-cleaning process is carried out followed by a conventional silicide process. See Fig. 9 and silicide 23. Standard back-end-of-line processing is done including pre-metal dielectric deposition and planarization, contact etch, contact liner formation and contact formation, followed by metal wiring and final chip fabrication, all not shown.

[0024] Simulation results indicate that the tensile stress level in uncapped polySi increases by about 600MPa to about 1200MPa as a result of annealing for about one hour at a temperature of about 600° C, while the capped polySi increases in tensile stress by only about 10MPa. Our simulation results show that about 33% to about 50% of the stress in the gate material (with the opposite sign with respect to the gate stress) can be translated into the channel region. Thus, the uncapped gate stack imparts – 200MPa to –300MPa, while the capped gate stack translates little or no stress into the channel region. One simulation of stress contours is shown in the diagram of Fig. 10.

[0025] While there has been shown and described what is at

present considered the preferred embodiments of the present invention, it will be apparent to those skilled in the art that various changes and modifications may be made therein without departing from the spirit and scope of the present invention which shall be limited only by the appended claims.